Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L4	7649	heterologous adj sequence	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 07:53
L5	83	L4 and glp	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 07:53
L6	12	L5 and preproglu$	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 07:55
L7	229	preproglu$	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 07:55
L8	53	l7 and @py<"2001"	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 07:56
L9	39	l7 and @py<"2000"	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 07:56
L10	18	l7 and @py<"1998"	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 08:55
L11	6	"6110707" and linker	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 08:55
L12	10	"6110707"	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 08:56
L13	3140	"6110707" glp	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 08:56
L14	3	"6110707" and glp	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 08:59
L15	2	l14 and fusion adj protein	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 09:14
L16	0	connolley adj s	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 09:14
L17	0	connolley and glp	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 09:14
L18	0	connelley and glp	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 09:15

L19	12	connelly and glp	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 09:19
L20	6	l19 and fusion adj protein	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 09:21
L21	2	"6723530"	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 09:21
L22	1	l21 and glp	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 10:10
L23	326	prepro$ and glp	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 10:11
L24	37	l23 and @py<"2000"	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 10:11
L25	10	l24 and heterologous	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 10:14
L26	599	glp and fusion adj protein	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 10:15
L27	40	l26 and @py<"2000"	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 10:15
L28	13	l26 and @py<"1998"	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 10:20
L29	72	l26 and @py>"1998" and @py<"2002"	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 10:21
L30	15	l26 and @py>"1998" and @py<"2000"	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 11:16
L31	18	wadsworth adj samuel	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 11:16
L32	20	armentano adj donna	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 11:16
L33	120	gregory adj richard	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 11:17
L34	3	parsons adj geoffrey	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 11:17

L35	137	l31 or l32 or l33 or l34	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 11:17
L36	3	l35 and glp	US-PGPUB; USPAT; DERWENT	OR	ON	2006/01/27 11:18

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 3, 2006, 23:27:35 ; Search time 229 Seconds
(without alignments)
240.311 Million cell updates/sec

Title: US-10-716-326-4
Perfect score: 402
Sequence: 1 MKIILWLCVFGLFLATLFPI.....SYLEGQAAKEFIAWLVKGRG 78

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	324	80.6	87	1	EXE4_HELSU	P26349 heloderma s
2	314	78.1	87	2	Q7SZU6_HELHR	Q7szu6 heloderma h
3	168.5	41.9	176	1	GLUC_SHEEP	Q8mj25 o glucagon
4	168.5	41.9	180	1	GLUC_BOVIN	P01272 b glucagon
5	168.5	41.9	180	1	GLUC_CANFA	P29794 c glucagon
6	168.5	41.9	180	1	GLUC_PIG	P01274 s glucagon
7	167.5	41.7	180	1	GLUC_CAVPO	P05110 c glucagon
8	167.5	41.7	180	1	GLUC_HUMAN	P01275 h glucagon
9	167.5	41.7	180	1	GLUC_MESAU	P01273 m glucagon
10	167.5	41.7	180	1	GLUC_MOUSE	P55095 m glucagon
11	167.5	41.7	180	1	GLUC_RAT	P06883 r glucagon
12	167.5	41.7	180	2	Q53TP6_HUMAN	Q53tp6 homo sapien
13	167	41.5	180	1	GLUC_OCTDE	P22890 o glucagon
14	166.5	41.4	45	2	Q6PPF4_CAPHI	Q6ppf4 capra hircu
15	164.5	40.9	206	1	GLUC_CHICK	P68259 g glucagon

16	154.5	38.4	124	2	Q6RYB1_9SAUR	Q6ryb1	agkistrodon
17	153.5	38.2	120	2	Q6RYB7_ICTPU	Q6ryb7	ictalurus p
18	151.5	37.7	204	1	GLUC_HELSU	O12956	h glucagon
19	149	37.1	145	2	Q6RYB5_NEOFS	Q6ryb5	neoceratodu
20	148	36.8	153	2	Q6RYB6_PRODO	Q6ryb6	protopterus
21	146.5	36.4	122	1	GLUC2_LOPAM	P04092	lophius ame
22	144	35.8	80	2	Q6IUP8_PHOSU	Q6iup8	phodopus su
23	140	34.8	123	2	Q6RYA9_9PERC	Q6rya9	sebastes ca
24	139	34.6	860	2	Q4RQJ4_TETNG	Q4rqj4	tetraodon n
25	138	34.3	220	2	Q8UWL9_9NEOB	Q8uwl9	hoplobatrac
26	134	33.3	149	2	Q6RYB2_BUFMA	Q6ryb2	bufo marinu
27	130	32.3	266	1	GLUC1_XENLA	O42143	xenopus lae
28	130	32.3	266	2	Q6DIZ4_XENTR	Q6diz4	xenopus tro
29	128	31.8	121	2	Q5PR39_BRARE	Q5pr39	brachydanio
30	127	31.6	103	1	GLUC_RANCA	P15438	rana catesb
31	125	31.1	121	2	Q6RYC1_9PERC	Q6ryc1	sebastes ca
32	125	31.1	176	2	Q6RYC2_9PERC	Q6ryc2	sebastes ca
33	125	31.1	219	1	GLUC2_XENLA	O42144	xenopus lae
34	125	31.1	219	2	Q5D082_XENLA	Q5d082	xenopus lae
35	124	30.8	66	2	Q788W6_ONCTS	Q788w6	oncorhynchu
36	124	30.8	72	2	Q91409_ONCTS	Q91409	oncorhynchu
37	124	30.8	122	2	Q6RYB8_ICTPU	Q6ryb8	ictalurus p
38	124	30.8	173	2	Q6RYB9_ICTPU	Q6ryb9	ictalurus p
39	124	30.8	178	1	GLUC1_ONCMY	Q91971	oncorhynchu
40	122.5	30.5	176	2	Q6RYB0_9PERC	Q6ryb0	sebastes ca
41	122	30.3	30	1	GLUCL_ANGAN	P63294	anguilla an
42	122	30.3	30	1	GLUCL_ANGRO	P63295	anguilla ro
43	120	29.9	121	1	GLUC_CARAU	P79695	carassius a
44	120	29.9	124	2	Q4S308_TETNG	Q4s308	tetraodon n
45	119	29.6	160	1	GLUC1_PETMA	Q9pur1	petromyzon

ALIGNMENTS

```
RESULT 1
EXE4_HELSU
ID   EXE4_HELSU     STANDARD;      PRT;    87 AA.
AC   P26349;
DT   01-MAY-1992 (Rel. 22, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Exendin-4 precursor.
OS   Heloderma suspectum (Gila monster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Helodermatidae;
OC   Heloderma.
OX   NCBI_TaxID=8554;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=97172477; PubMed=9020121; DOI=10.1074/jbc.272.7.4335;
RA   Chen Y.E., Drucker D.J.;
RT   "Tissue-specific expression of unique mRNAs that encode proglucagon-
RT   derived peptides or exendin 4 in the lizard.";
RL   J. Biol. Chem. 272:4108-4115(1997).
RN   [2]
RP   PROTEIN SEQUENCE OF 48-86.
```

```
RC   TISSUE=Venom;
RX   MEDLINE=92218391; PubMed=1313797;
RA   Eng J., Kleinman W.A., Singh L., Singh G., Raufman J.-P.;
RT   "Isolation and characterization of exendin-4, an exendin-3 analogue,
RT   from Heloderma suspectum venom. Further evidence for an exendin
RT   receptor on dispersed acini from guinea pig pancreas.";
RL   J. Biol. Chem. 267:7402-7405(1992).
RN   [3]
RP   STRUCTURE BY NMR OF 48-86.
RX   PubMed=11683627; DOI=10.1021/bi010902s;
RA   Neidigh J.W., Fesinmeyer R.M., Prickett K.S., Andersen N.H.;
RT   "Exendin-4 and glucagon-like-peptide-1: NMR structural comparisons in
RT   the solution and micelle-associated states.";
RL   Biochemistry 40:13188-13200(2001).
CC   -!- FUNCTION: Has a VIP/secretin-like biological activity. Interacts
CC       with the exendin receptor.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: Belongs to the glucagon family.
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; U77613; AAB51130.1; -; mRNA.
DR   PIR; A42486; HWGH4G.
DR   PDB; 1JRJ; NMR; A=48-86.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; Hormone_2; 1.
DR   SMART; SM00070; GLUCA; 1.
DR   PROSITE; PS00260; GLUCAGON; 1.
KW   3D-structure; Amidation; Direct protein sequencing; Glucagon family;
KW   Signal; Toxin.
FT   SIGNAL        1     23       Potential.
FT   PROPEP       24     47
FT   PEPTIDE      48     86       Exendin-4.
FT   MOD_RES      86     86       Serine amide (G-87 provides amide group).
FT   TURN         52     53
FT   HELIX        54     74
FT   TURN         75     76
FT   HELIX        77     79
SQ   SEQUENCE   87 AA;  9479 MW;  656BA6E3D87454A2 CRC64;

  Query Match             80.6%;  Score 324;  DB 1;  Length 87;
  Best Local Similarity   81.6%;  Pred. No. 5.4e-28;
  Matches   62;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 MKIILWLCVFGLFLATLFPISWQMPVESGLSSEDSASSESFAKRIKRHGEGTFTSDVSSY 60
              |||||||||||||||||||||||||||||||||||||||| :|||||||||||:|
Db          1 MKIILWLCVFGLFLATLFPISWQMPVESGLSSEDSASSESFASKIKRHGEGTFTSDLSKQ 60

Qy         61 LEGQAAKEFIAWLVKG 76
              :| :| : || ||  |
Db         61 MEEEAVRLFIEWLKNG 76
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 3, 2006, 23:17:55 ; Search time 185 Seconds
(without alignments)
185.252 Million cell updates/sec

Title: US-10-716-326-4
Perfect score: 402
Sequence: 1 MKIILWLCVFGLFLATLFPI.....SYLEGQAAKEFIAWLVKGRG 78

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	402	100.0	78	6	ABP98121	Abp98121 Amino aci
2	402	100.0	78	9	ADZ80126	Adz80126 Exendin-4
3	364	90.5	80	6	ABP98129	Abp98129 Amino aci
4	364	90.5	80	9	ADZ80142	Adz80142 Exendin-4
5	325.5	81.0	77	6	ABP98122	Abp98122 Amino aci
6	325.5	81.0	77	9	ADZ80128	Adz80128 Helodermi
7	324	80.6	87	2	AAW70288	Aaw70288 Heloderma
8	324	80.6	87	7	ADF15336	Adf15336 Human alb

9	324	80.6	87	7	ADF15337	Adf15337 Human alb
10	324	80.6	87	7	ADH21405	Adh21405 Human ext
11	324	80.6	87	7	ADH21406	Adh21406 Human ext
12	231	57.5	47	6	ABU91976	Abu91976 Glia mons
13	198	49.3	64	8	ADL92153	Adl92153 Exendin-4
14	195	48.5	84	6	ABP98124	Abp98124 Amino aci
15	195	48.5	84	9	ADZ80132	Adz80132 IGF-1(fur
16	192	47.8	82	6	ABP98123	Abp98123 Amino aci
17	192	47.8	82	9	ADZ80130	Adz80130 GIP.GLP-1
18	184.5	45.9	77	6	ABP98128	Abp98128 Amino aci
19	184.5	45.9	77	9	ADZ80140	Adz80140 Factor IX
20	178	44.3	386	4	AAB30703	Aab30703 A Bacillu
21	176.5	43.9	77	7	ADF16689	Adf16689 Human alb
22	176.5	43.9	77	7	ADH21889	Adh21889 Human GLP
23	176.5	43.9	662	7	ADF16526	Adf16526 Human alb
24	176.5	43.9	662	7	ADH21814	Adh21814 Human alb
25	176	43.8	654	7	ADF16462	Adf16462 Human alb
26	176	43.8	654	7	ADH21777	Adh21777 Human alb
27	172.5	42.9	89	7	ADF16195	Adf16195 Human alb
28	172.5	42.9	89	7	ADH21652	Adh21652 Human GLP
29	172.5	42.9	673	7	ADF17044	Adf17044 Human alb
30	172.5	42.9	673	7	ADH22037	Adh22037 Mouse alb
31	172.5	42.9	674	7	ADF16193	Adf16193 Human alb
32	172.5	42.9	674	7	ADH21650	Adh21650 Human alb
33	172.5	42.9	674	9	ADW45202	Adw45202 K. lactis
34	172.5	42.9	915	9	ADW45204	Adw45204 K. lactis
35	172	42.8	663	7	ADF16512	Adf16512 Human alb
36	172	42.8	663	7	ADH21803	Adh21803 Human alb
37	171.5	42.7	60	7	ADF17048	Adf17048 Human alb
38	171.5	42.7	60	7	ADH22038	Adh22038 Human GLP
39	171.5	42.7	83	7	ADF16687	Adf16687 Human alb
40	171.5	42.7	83	7	ADH21887	Adh21887 Human GLP
41	171.5	42.7	145	7	ADF16688	Adf16688 Human alb
42	171.5	42.7	145	7	ADH21888	Adh21888 Human GLP
43	171.5	42.7	664	7	ADF16510	Adf16510 Human alb
44	171.5	42.7	664	7	ADH21801	Adh21801 Human alb
45	171.5	42.7	668	7	ADF16524	Adf16524 Human alb

ALIGNMENTS

```
RESULT 1
ABP98121
ID   ABP98121 standard; protein; 78 AA.
XX
AC   ABP98121;
XX
DT   11-AUG-2003  (first entry)
XX
DE   Amino acid sequence of proexendin leader linked to modified GLP-1.
XX
KW   Glucagon-like peptide 1; GLP-1; blood sugar disorder; diabetes;
KW   hyperglycemia; hypoglycaemia; hypoinsulinism; insulin production;
KW   proexendin.
XX
OS   Synthetic.
```

OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1. .47
FT /note= "proexendin leader"
XX
PN WO2003014318-A2.
XX
PD 20-FEB-2003.
XX
PF 07-AUG-2002; 2002WO-US025227.
XX
PR 08-AUG-2001; 2001US-0310982P.
XX
PA (GENZ) GENZYME CORP.
XX
PI Wadsworth SC, Armentano D, Gregory RJ, Parsons G;
XX
DR WPI; 2003-256570/25.
DR N-PSDB; ACC43465.
XX
PT New glucagon-like peptide 1 (GLP-1) precursor comprising a mammalian GLP-
PT 1 linked to a heterologous signal sequence, useful for treating blood
PT sugar disorders, e.g. Type I and Type II diabetes, hyperglycemia, or
PT hypoglycemia.
XX
PS Claim 16; Fig 2; 69pp; English.
XX
CC The present sequence is a proexendin leader linked to a Gly-8 modified
CC human glucagon-like peptide-1 (GLP-1). The specification describes
CC compositions, expression vectors and host cells comprising nucleic acid
CC which encodes a precursor GLP-1 linked to a heterologous signal sequence.
CC Nucleic acids encoding GLP-1 may be used to treat blood sugar disorders
CC such as Type I and Type II diabetes, hyperglycemia, hypoglycaemia, and
CC hypoinsulinism and to promote or stimulate insulin production
XX
SQ Sequence 78 AA;

Query Match 100.0%; Score 402; DB 6; Length 78;
Best Local Similarity 100.0%; Pred. No. 2.4e-40;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKIILWLCVFGLFLATLFPISWQMPVESGLSSEDSASSESFAKRIKRHGEGTFTSDVSSY 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKIILWLCVFGLFLATLFPISWQMPVESGLSSEDSASSESFAKRIKRHGEGTFTSDVSSY 60

Qy         61 LEGQAAKEFIAWLVKGRG 78
              ||||||||||||||||||
Db         61 LEGQAAKEFIAWLVKGRG 78
```

RESULT 2
ADZ80126
ID ADZ80126 standard; protein; 78 AA.
XX
AC ADZ80126;
XX

DT 28-JUL-2005 (first entry)
XX
DE Exendin-4.GLP-1Gly8 amino acid sequence SEQ ID NO:4.
XX
KW glucagon-like peptide 1; fusion protein; antidiabetic; hepatotropic;
KW metabolic; insulin dependent diabetes; non-insulin dependent diabetes;
KW metabolic disorder; liver disease; exendin.
XX
OS Mammalia.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1. .47
FT /note= "Proexendin signal peptide"
FT Protein 48. .78
FT /note= "GLP-1Gly8"
XX
PN US2005107318-A1.
XX
PD 19-MAY-2005.
XX
PF 17-NOV-2003; 2003US-00715976.
XX
PR 17-NOV-2003; 2003US-00715976.
XX
PA (WADS/) WADSWORTH S.
PA (ARME/) ARMENTANO D.
PA (GREG/) GREGORY R J.
PA (PARS/) PARSONS G.
XX
PI Wadsworth S, Armentano D, Gregory RJ, Parsons G;
XX
DR WPI; 2005-354819/36.
DR N-PSDB; ADZ80125.
XX
PT Treating diabetes comprises administering a nucleic acid encoding a
PT precursor Glucagon-like peptide 1 (GLP-1) comprising mammalian GLP-1
PT linked to a heterologous signal sequence.
XX
PS Example 1; SEQ ID NO 4; 70pp; English.
XX
CC The invention relates to a method which comprises administering to an
CC individual a nucleic acid encoding a precursor glucagon-like peptide 1
CC (GLP-1) comprising mammalian GLP-1 linked to a heterologous signal
CC sequence, where the precursor GLP-1 is cleaved in vivo or ex vivo which
CC results in generation of activated GLP-1 in the individual, where the
CC activated GLP-1 reduces plasma triglyceride levels or lipid accumulation
CC in an organ. Also described: (1) nucleic acids which encode GLP-1; (2)
CC vectors comprising the nucleic acids; and (3) host cells comprising the
CC nucleic acid. The method is useful for treating Type I diabetes (insulin
CC dependent diabetes) and Type II diabetes (non-insulin dependent
CC diabetes), metabolic syndrome and non-alcoholic fatty liver disease. The
CC present sequence represents the exendin-4.GLP-1Gly8 GLP-1 fusion protein,
CC which is used in an example from the present invention.
XX
SQ Sequence 78 AA;

```
  Query Match             100.0%;  Score 402;  DB 9;  Length 78;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-40;
  Matches   78;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKIILWLCVFGLFLATLFPISWQMPVESGLSSEDSASSESFAKRIKRHGEGTFTSDVSSY 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKIILWLCVFGLFLATLFPISWQMPVESGLSSEDSASSESFAKRIKRHGEGTFTSDVSSY 60

Qy         61 LEGQAAKEFIAWLVKGRG 78
              ||||||||||||||||||
Db         61 LEGQAAKEFIAWLVKGRG 78


RESULT 3
ABP98129
ID   ABP98129 standard; protein; 80 AA.
XX
AC   ABP98129;
XX
DT   11-AUG-2003  (first entry)
XX
DE   Amino acid sequence of proexendin leader linked to modified GLP-1.
XX
KW   Glucagon-like peptide 1; GLP-1; blood sugar disorder; diabetes;
KW   hyperglycemia; hypoglycaemia; hypoinsulinism; insulin production;
KW   proexendin.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .40
FT                   /note= "proexendin leader"
XX
PN   WO2003014318-A2.
XX
PD   20-FEB-2003.
XX
PF   07-AUG-2002; 2002WO-US025227.
XX
PR   08-AUG-2001; 2001US-0310982P.
XX
PA   (GENZ ) GENZYME CORP.
XX
PI   Wadsworth SC,  Armentano D,  Gregory RJ,  Parsons G;
XX
DR   WPI; 2003-256570/25.
DR   N-PSDB; ACC43473.
XX
PT   New glucagon-like peptide 1 (GLP-1) precursor comprising a mammalian GLP-
PT   1 linked to a heterologous signal sequence, useful for treating blood
PT   sugar disorders, e.g. Type I and Type II diabetes, hyperglycemia, or
PT   hypoglycemia.
XX
PS   Claim 16; Fig 10; 69pp; English.
XX
CC   The present sequence represents a proexendin leader linked to a Gly-8
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 5, 2006, 00:00:23 ; Search time 169 Seconds
(without alignments)
820.413 Million cell updates/sec

Title: US-10-716-326-4
Perfect score: 402
Sequence: 1 MKIILWLCVFGLFLATLFPI.....SYLEGQAAKEFIAWLVKGRG 78

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-
Q=/cgn2_1/USPTO_spool/US10716326/runat_03012006_143904_8663/app_query.fasta_1.26
3
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10716326_@CGN_1_1_193_@runat_03012006_143904_8663 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		324	80.6	492	3	US-09-019-172-1	Sequence 1, Appli
2		167.5	41.7	895	3	US-08-784-582-55	Sequence 55, Appl
3		167.5	41.7	955	3	US-08-784-582-57	Sequence 57, Appl
4		167.5	41.7	955	3	US-08-784-582-60	Sequence 60, Appl
5		167.5	41.7	1034	3	US-09-635-679E-1	Sequence 1, Appli
6		167.5	41.7	1062	3	US-09-016-434-1212	Sequence 1212, Ap
7		167.5	41.7	2356	3	US-08-784-582-72	Sequence 72, Appl
8		163	40.5	144	2	US-08-835-231-17	Sequence 17, Appl
9		163	40.5	144	3	US-09-108-661-17	Sequence 17, Appl
10		157	39.1	528	2	US-08-835-231-7	Sequence 7, Appli
11		157	39.1	528	2	US-08-835-231-8	Sequence 8, Appli
12		157	39.1	528	3	US-09-108-661-7	Sequence 7, Appli
13		157	39.1	528	3	US-09-108-661-8	Sequence 8, Appli
14		118	29.4	78	2	US-08-829-876-22	Sequence 22, Appl
15		118	29.4	78	3	US-09-234-874A-22	Sequence 22, Appl
16		118	29.4	78	3	US-09-234-873A-22	Sequence 22, Appl
17		118	29.4	78	3	US-10-076-604-22	Sequence 22, Appl
18		102	25.4	207	2	US-08-193-863-3	Sequence 3, Appli
19		102	25.4	207	2	US-08-377-833-3	Sequence 3, Appli
20		102	25.4	207	2	US-08-324-502-3	Sequence 3, Appli
21		102	25.4	207	2	US-08-083-501-3	Sequence 3, Appli
22		102	25.4	207	2	US-08-415-939-3	Sequence 3, Appli
23		99	24.6	110	2	US-07-741-931-8	Sequence 8, Appli
c 24		99	24.6	110	2	US-07-741-931-9	Sequence 9, Appli
25		99	24.6	110	2	US-07-937-132A-8	Sequence 8, Appli
c 26		99	24.6	110	2	US-07-937-132A-9	Sequence 9, Appli
27		93	23.1	138	3	US-09-614-847-146	Sequence 146, App
28		89.5	22.3	255	3	US-09-280-030-51	Sequence 51, Appl
29		89	22.1	57	2	US-08-811-028-43	Sequence 43, Appl
30		89	22.1	87	2	US-07-741-931-1	Sequence 1, Appli
31		89	22.1	87	2	US-07-937-132A-1	Sequence 1, Appli
32		89	22.1	87	3	US-09-280-030-59	Sequence 59, Appl
c 33		85	21.1	73	2	US-08-829-876-24	Sequence 24, Appl
c 34		85	21.1	73	3	US-09-234-874A-24	Sequence 24, Appl
c 35		85	21.1	73	3	US-09-234-873A-24	Sequence 24, Appl
c 36		85	21.1	73	3	US-10-076-604-24	Sequence 24, Appl
37		84	20.9	107	2	US-07-741-931-10	Sequence 10, Appl
38		84	20.9	107	2	US-07-937-132A-10	Sequence 10, Appl
c 39		79	19.7	48	2	US-08-811-028-44	Sequence 44, Appl
c 40		79	19.7	107	2	US-07-741-931-11	Sequence 11, Appl
c 41		79	19.7	107	2	US-07-937-132A-11	Sequence 11, Appl
c 42		78	19.4	53	2	US-08-811-028-46	Sequence 46, Appl
43		74.5	18.5	42954	3	US-09-949-016-17123	Sequence 17123, A
44		74.5	18.5	42954	3	US-09-949-016-17124	Sequence 17124, A
45		71.5	17.8	711	3	US-09-949-016-3581	Sequence 3581, Ap

ALIGNMENTS

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 4, 2006, 17:16:20 ; Search time 3647 Seconds
(without alignments)
1215.736 Million cell updates/sec

Title: US-10-716-326-4
Perfect score: 402
Sequence: 1 MKIILWLCVFGLFLATLFPI.....SYLEGQAAKEFIAWLVKGRG 78

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-
Q=/cgn2_1/USPTO_spool/US10716326/runat_03012006_143904_8645/app_query.fasta_1.26
3
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10716326_@CGN_1_1_4939_@runat_03012006_143904_8645 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*

11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	324	80.6	465	5	HSU77613	U77613 Heloderma s
2	324	80.6	492	6	AR526929	AR526929 Sequence
3	314	78.1	397	5	HHO580309	AJ580309 Heloderma
4	314	78.1	471	6	AX824524	AX824524 Sequence
5	187.5	46.6	458	6	AX824523	AX824523 Sequence
6	172	42.8	396	6	AX147675	AX147675 Sequence
7	171.5	42.7	1104	9	OCOGLU	M57688 Octodon deg
8	168.5	41.9	559	4	AF529185	AF529185 Ovis arie
9	168.5	41.9	574	4	AY242124	AY242124 Sus scrof
10	168.5	41.9	1108	4	BOVGG	K00107 Bovine panc
11	168.5	41.9	1123	4	AF308439	AF308439 Canis fam
12	167.5	41.7	543	6	CS077314	CS077314 Sequence
13	167.5	41.7	543	8	BT006813	BT006813 Homo sapi
14	167.5	41.7	543	11	AY890068	AY890068 Synthetic
15	167.5	41.7	543	11	AY890069	AY890069 Synthetic
16	167.5	41.7	543	11	AY892547	AY892547 Synthetic
17	167.5	41.7	543	11	AY892548	AY892548 Synthetic
18	167.5	41.7	543	11	BT007507	BT007507 Synthetic
19	167.5	41.7	668	9	MMPPROGLG	Z46845 M.musculus
20	167.5	41.7	895	6	AR108106	AR108106 Sequence
21	167.5	41.7	955	6	AR108107	AR108107 Sequence
22	167.5	41.7	955	6	AR108109	AR108109 Sequence
23	167.5	41.7	1034	6	A31421	A31421 H.sapiens m
24	167.5	41.7	1034	6	AR634243	AR634243 Sequence
25	167.5	41.7	1036	6	CQ729100	CQ729100 Sequence
26	167.5	41.7	1053	9	GPIGG	D00014 Cavia porce
27	167.5	41.7	1062	6	AR270649	AR270649 Sequence
28	167.5	41.7	1062	8	HUMGLUC	J04040 Human gluca
29	167.5	41.7	1102	9	BC012975	BC012975 Mus muscu
30	167.5	41.7	1116	9	AF276754	AF276754 Mus muscu
31	167.5	41.7	1118	9	HAMGG	J00059 Syrian hams
32	167.5	41.7	1128	6	CS077369	CS077369 Sequence
33	167.5	41.7	1154	8	BC005278	BC005278 Homo sapi
34	167.5	41.7	2356	6	AR108119	AR108119 Sequence
35	167.5	41.7	3798	6	CQ493428	CQ493428 Sequence
36	167.5	41.7	4146	6	CQ414344	CQ414344 Sequence
37	166.5	41.4	228	6	AX840858	AX840858 Sequence
38	166.5	41.4	277	9	RATGLU4	K02811 Rat glucago
39	166.5	41.4	707	4	AY588290	AY588290 Capra hir
40	166.5	41.4	6455	8	HSGLUC	V01515 Human gene
41	166.5	41.4	10050	8	HSGLUCG2	X03991 Human gluca
c 42	166.5	41.4	163681	8	AC007750	AC007750 Homo sapi

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 3, 2006, 23:30:36 ; Search time 38 Seconds
(without alignments)
197.498 Million cell updates/sec

Title: US-10-716-326-4
Perfect score: 402
Sequence: 1 MKIILWLCVFGLFLATLFPI.....SYLEGQAAKEFIAWLVKGRG 78

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	168.5	41.9	180	1	GCBO	glucagon precursor
2	167.5	41.7	180	1	GCHY	glucagon precursor
3	167.5	41.7	180	1	GCGP	glucagon precursor
4	167.5	41.7	180	1	GCHU	glucagon precursor
5	167.5	41.7	180	1	GCRT	glucagon precursor
6	167.5	41.7	180	2	A57294	glucagon precursor
7	167	41.5	180	1	GCRTDU	glucagon precursor
8	164.5	40.9	151	1	GCCH	glucagon precursor
9	164.5	40.9	206	2	I51301	proglucagon - chic
10	157.5	39.2	158	1	GCPG	glucagon precursor
11	146.5	36.4	122	1	GCAF2	glucagon 2 precurs
12	128	31.8	101	1	GCFGB	glucagon precursor
13	124	30.8	66	2	I51093	glucagon - chinook

14	124	30.8	178	2	I51058	glucagon I precurs
15	122	30.3	30	2	C61125	glucagon-like pept
16	122	30.3	30	2	B61125	glucagon-like pept
17	119	29.6	178	2	I51057	glucagon II precur
18	113	28.1	63	1	GCIDC	glucagon precursor
19	112	27.9	72	1	GCGXA	glucagon precursor
20	111.5	27.7	124	1	GCAF	glucagon 1 precurs
21	109	27.1	60	1	GCONC	glucagon precursor
22	107	26.6	30	2	S44473	glucagon-like pept
23	103.5	25.7	87	1	GCFIS	glucagon precursor
24	102	25.4	69	1	GCDG69	glucagon-69 - dog
25	96	23.9	29	2	S07211	glucagon - marbled
26	95	23.6	31	2	S44472	glucagon G2 - Nort
27	94	23.4	29	1	GCDF	glucagon - smaller
28	93	23.1	31	2	S44471	glucagon G1 - Nort
29	93	23.1	39	1	HWGH4G	exendin-4 - Gila m
30	92	22.9	29	1	GCEN	glucagon - elephan
31	89	22.1	29	1	GCOPV	glucagon - North A
32	89	22.1	29	2	A91740	glucagon - turkey
33	89	22.1	29	2	C39258	glucagon - common
34	89	22.1	29	2	A91742	glucagon - Arabian
35	89	22.1	29	2	A91741	glucagon - rabbit
36	87	21.6	29	1	A61583	glucagon - ostrich
37	87	21.6	29	1	GCDK	glucagon - duck
38	87	21.6	29	1	GCTTS	glucagon - slider
39	87	21.6	29	2	C60840	glucagon I - Europ
40	86	21.4	29	1	GCCB	glucagon - Chinchi
41	85	21.1	29	1	GCFLE	glucagon - Europea
42	85	21.1	29	2	A61135	glucagon - bigeye
43	84	20.9	39	1	HWGH3Z	exendin-3 - Mexica
44	82	20.4	29	2	S39018	glucagon - bowfin
45	80	19.9	144	2	S71426	glucose-dependent

ALIGNMENTS

RESULT 1
GCBO
glucagon precursor - bovine
N;Contains: glicentin-related peptide; glucagon; glucagon-like peptide 1;
glucagon-like peptide 2
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 20-Mar-1998
C;Accession: A93970; A92081; A01538
R;Lopez, L.C.; Frazier, M.L.; Su, C.J.; Kumar, A.; Saunders, G.F.
Proc. Natl. Acad. Sci. U.S.A. 80, 5485-5489, 1983
A;Title: Mammalian pancreatic preproglucagon contains three glucagon-related
peptides.
A;Reference number: A93970; MUID:83299996; PMID:6577439
A;Accession: A93970
A;Molecule type: mRNA
A;Residues: 1-180 <LOP>
A;Cross-references: UNIPARC:UPI00001734FF; EMBL:K00107
R;Bromer, W.W.; Boucher, M.E.; Koffenberger Jr., J.E.
J. Biol. Chem. 246, 2822-2827, 1971
A;Title: Amino acid sequence of bovine glucagon.

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 3, 2006, 23:33:51 ; Search time 12 Seconds
(without alignments)
48.677 Million cell updates/sec

Title: US-10-716-326-4
Perfect score: 402
Sequence: 1 MKIILWLCVFGLFLATLFPI.....SYLEGQAAKEFIAWLVKGRG 78

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 57103 seqs, 7488799 residues

Total number of hits satisfying chosen parameters: 57103

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	169	42.0	70	6	US-10-997-061-28	Sequence 28, Appl
2	161.5	40.2	117	6	US-10-997-061-9	Sequence 9, Appli
3	161.5	40.2	122	6	US-10-997-061-31	Sequence 31, Appl
4	158	39.3	277	6	US-10-997-061-11	Sequence 11, Appl
5	157.5	39.2	119	6	US-10-997-061-13	Sequence 13, Appl
6	157	39.1	31	6	US-10-997-061-3	Sequence 3, Appli
7	154	38.3	31	6	US-10-997-061-7	Sequence 7, Appli
8	151	37.6	30	6	US-10-997-061-1	Sequence 1, Appli
9	151	37.6	31	6	US-10-997-061-4	Sequence 4, Appli

10	151	37.6	31	7	US-11-112-277-9	Sequence 9, Appli
11	151	37.6	31	7	US-11-112-277-23	Sequence 23, Appl
12	151	37.6	31	7	US-11-112-277-24	Sequence 24, Appl
13	151	37.6	34	6	US-10-997-061-10	Sequence 10, Appl
14	151	37.6	34	6	US-10-997-061-12	Sequence 12, Appl
15	150	37.3	31	7	US-11-112-277-1	Sequence 1, Appli
16	150	37.3	31	7	US-11-112-277-10	Sequence 10, Appl
17	148	36.8	30	6	US-10-997-061-5	Sequence 5, Appli
18	148	36.8	31	6	US-10-997-061-8	Sequence 8, Appli
19	147	36.6	30	7	US-11-112-277-20	Sequence 20, Appl
20	146	36.3	30	6	US-10-997-061-2	Sequence 2, Appli
21	145	36.1	30	7	US-11-112-277-11	Sequence 11, Appl
22	145	36.1	30	7	US-11-112-277-12	Sequence 12, Appl
23	145	36.1	30	7	US-11-112-277-19	Sequence 19, Appl
24	145	36.1	30	7	US-11-112-277-21	Sequence 21, Appl
25	145	36.1	30	7	US-11-112-277-22	Sequence 22, Appl
26	145	36.1	31	7	US-11-112-277-26	Sequence 26, Appl
27	143	35.6	29	7	US-11-112-277-32	Sequence 32, Appl
28	143	35.6	30	6	US-10-997-061-6	Sequence 6, Appli
29	98	24.4	39	7	US-11-174-089-225	Sequence 225, App
30	97	24.1	39	7	US-11-174-089-222	Sequence 222, App
31	95	23.6	39	7	US-11-174-089-227	Sequence 227, App
32	93	23.1	29	7	US-11-174-089-66	Sequence 66, Appl
33	93	23.1	30	7	US-11-174-089-3	Sequence 3, Appli
34	93	23.1	30	7	US-11-174-089-4	Sequence 4, Appli
35	93	23.1	30	7	US-11-174-089-26	Sequence 26, Appl
36	93	23.1	31	7	US-11-174-089-63	Sequence 63, Appl
37	93	23.1	32	7	US-11-174-089-61	Sequence 61, Appl
38	93	23.1	33	7	US-11-174-089-59	Sequence 59, Appl
39	93	23.1	34	7	US-11-174-089-57	Sequence 57, Appl
40	93	23.1	35	7	US-11-174-089-55	Sequence 55, Appl
41	93	23.1	36	7	US-11-174-089-53	Sequence 53, Appl
42	93	23.1	36	7	US-11-174-089-73	Sequence 73, Appl
43	93	23.1	37	7	US-11-174-089-51	Sequence 51, Appl
44	93	23.1	37	7	US-11-174-089-70	Sequence 70, Appl
45	93	23.1	37	7	US-11-174-089-71	Sequence 71, Appl

ALIGNMENTS

RESULT 1
US-10-997-061-28
; Sequence 28, Application US/10997061
; Publication No. US20050260701A1
; GENERAL INFORMATION:
; APPLICANT: Wagner, Fred W.
; APPLICANT: Luan, Peng
; APPLICANT: Xia, Yuannan
; APPLICANT: Bossard, Mary
; APPLICANT: Holmquist, Barton
; APPLICANT: Merrifield, Edwin H.
; APPLICANT: Strydom, Daniel
; APPLICANT: Restoragen Inc.
; TITLE OF INVENTION: Method for Enzymatic Production of GLP-1(7-36) Amide Peptides
; FILE REFERENCE: 1627.003US1

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 3, 2006, 23:23:10 ; Search time 47 Seconds
(without alignments)
137.206 Million cell updates/sec

Title: US-10-716-326-4
Perfect score: 402
Sequence: 1 MKIILWLCVFGLFLATLFPI.....SYLEGQAAKEFIAWLVKGRG 78

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	324	80.6	87	2	US-09-019-172-2	Sequence 2, Appli
2	198	49.3	64	2	US-10-360-101-228	Sequence 228, App
3	168	41.8	36	2	US-09-614-847-119	Sequence 119, App
4	168	41.8	42	2	US-09-614-847-118	Sequence 118, App
5	167.5	41.7	180	2	US-08-784-582-56	Sequence 56, Appl
6	167.5	41.7	180	2	US-08-784-582-58	Sequence 58, Appl
7	167.5	41.7	180	2	US-08-784-582-61	Sequence 61, Appl
8	167.5	41.7	180	2	US-09-635-679E-2	Sequence 2, Appli
9	167.5	41.7	360	2	US-08-784-582-73	Sequence 73, Appl
10	163	40.5	31	2	US-09-209-799D-16	Sequence 16, Appl
11	163	40.5	31	2	US-09-614-847-123	Sequence 123, App

12	163	40.5	31	2	US-09-997-792A-14	Sequence 14, Appl
13	163	40.5	32	2	US-09-614-847-147	Sequence 147, App
14	163	40.5	37	1	US-08-095-162-19	Sequence 19, Appl
15	163	40.5	37	1	US-08-470-220A-19	Sequence 19, Appl
16	163	40.5	37	2	US-08-967-374-19	Sequence 19, Appl
17	163	40.5	37	2	US-09-302-596-1	Sequence 1, Appli
18	163	40.5	37	2	US-08-472-349-1	Sequence 1, Appli
19	163	40.5	37	2	US-09-623-618B-1	Sequence 1, Appli
20	163	40.5	37	2	US-09-333-415-1	Sequence 1, Appli
21	163	40.5	37	2	US-09-505-991-19	Sequence 19, Appl
22	163	40.5	37	2	US-09-303-016-1	Sequence 1, Appli
23	163	40.5	37	2	US-09-657-332A-1	Sequence 1, Appli
24	163	40.5	37	2	US-09-614-847-122	Sequence 122, App
25	163	40.5	37	2	US-09-805-507-1	Sequence 1, Appli
26	163	40.5	37	2	US-09-876-388-1	Sequence 1, Appli
27	163	40.5	37	2	US-09-859-804-1	Sequence 1, Appli
28	163	40.5	37	2	US-10-287-892-1	Sequence 1, Appli
29	163	40.5	37	2	US-09-943-084-1	Sequence 1, Appli
30	163	40.5	37	2	US-09-623-548A-343	Sequence 343, App
31	163	40.5	37	2	US-10-055-259-1	Sequence 1, Appli
32	163	40.5	37	2	US-09-657-276-343	Sequence 343, App
33	163	40.5	37	2	US-10-288-340-1	Sequence 1, Appli
34	163	40.5	37	2	US-09-982-978-1	Sequence 1, Appli
35	160	39.8	31	2	US-09-258-750-14	Sequence 14, Appl
36	160	39.8	31	2	US-09-258-750-15	Sequence 15, Appl
37	160	39.8	31	2	US-09-258-750-16	Sequence 16, Appl
38	160	39.8	31	2	US-09-209-799D-28	Sequence 28, Appl
39	160	39.8	31	2	US-09-398-111-14	Sequence 14, Appl
40	160	39.8	31	2	US-09-398-111-15	Sequence 15, Appl
41	160	39.8	31	2	US-09-398-111-16	Sequence 16, Appl
42	160	39.8	31	2	US-09-997-792A-25	Sequence 25, Appl
43	160	39.8	32	2	US-09-258-750-95	Sequence 95, Appl
44	160	39.8	32	2	US-09-398-111-95	Sequence 95, Appl
45	160	39.8	33	2	US-09-258-750-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1
US-09-019-172-2
; Sequence 2, Application US/09019172A
; Patent No. 6723530
; GENERAL INFORMATION:
; APPLICANT: Drucker, Daniel J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING PROEXENDIN, AND METHODS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 8607-013
; CURRENT APPLICATION NUMBER: US/09/019,172A
; CURRENT FILING DATE: 1998-02-05
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 87
;   TYPE: PRT
;   ORGANISM: LIZARD
US-09-019-172-2

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 4, 2006, 21:56:58 ; Search time 3041 Seconds
(without alignments)
1200.063 Million cell updates/sec

Title: US-10-716-326-4
Perfect score: 402
Sequence: 1 MKIILWLCVFGLFLATLFPI.....SYLEGQAAKEFIAWLVKGRG 78

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-
Q=/cgn2_1/USPTO_spool/US10716326/runat_03012006_143904_8655/app_query.fasta_1.26
3
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10716326_@CGN_1_1_5315_@runat_03012006_143904_8655 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*

11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	178	44.3	557	2	BG656218	BG656218 ib38e07.y
	2	174.5	43.4	400	6	CB068992	CB068992 is33f09.y
	3	174.5	43.4	611	6	CA847981	CA847981 iq41a09.y
	4	173.5	43.2	509	3	BM315096	BM315096 ig43d08.y
	5	172.5	42.9	535	5	BU951082	BU951082 io75f03.y
	6	172.5	42.9	757	6	CA775068	CA775068 ip12f09.y
	7	172.5	42.9	928	2	BG116189	BG116189 602318388
	8	171	42.5	592	5	BU079338	BU079338 in22e10.y
	9	170.5	42.4	487	5	BQ787033	BQ787033 il49c01.y
	10	170.5	42.4	661	6	CA776177	CA776177 ip01g06.y
	11	169.5	42.2	494	2	BG311458	BG311458 ib25e03.y
	12	169.5	42.2	550	2	BG656010	BG656010 ib35h06.y
	13	169.5	42.2	604	5	BU078943	BU078943 im68e01.y
	14	169.5	42.2	648	6	CA775681	CA775681 io88h06.y
	15	169	42.0	508	2	BG655984	BG655984 ib35e10.y
	16	169	42.0	656	6	CA941719	CA941719 ir35b11.y
	17	168.5	41.9	425	8	DN351146	DN351146 LIB3579-0
	18	168.5	41.9	538	6	CA778313	CA778313 ip16h06.y
	19	168.5	41.9	576	3	BI774144	BI774144 466139 MA
	20	168.5	41.9	586	5	BX669011	BX669011 BX669011
	21	168.5	41.9	596	8	DN358389	DN358389 LIB3626-0
	22	168.5	41.9	603	6	CA778273	CA778273 ip16d02.y
	23	168.5	41.9	622	3	BM311910	BM311910 ig63f12.y
	24	168.5	41.9	627	8	DN357363	DN357363 LIB3626-0
	25	168.5	41.9	630	8	DN357084	DN357084 LIB3626-0
	26	168.5	41.9	651	6	CF788378	CF788378 860325 MA
	27	168.5	41.9	662	1	AW583897	AW583897 ia06d04.y
	28	168.5	41.9	680	8	DN356274	DN356274 LIB3626-0
	29	168.5	41.9	685	7	CK835133	CK835133 4059830 B
	30	168.5	41.9	697	7	CK951493	CK951493 4091010 B
	31	168.5	41.9	705	7	CN789364	CN789364 4123822 B
	32	168.5	41.9	709	7	CN792256	CN792256 4127095 B
	33	168.5	41.9	709	7	CN793557	CN793557 4128597 B
	34	168.5	41.9	713	7	CK832710	CK832710 4056519 B
	35	168.5	41.9	738	6	CA774895	CA774895 ip10b09.y
c	36	168.5	41.9	804	7	CK772000	CK772000 960421 MA
	37	168.5	41.9	837	8	DN516026	DN516026 1254447 M
	38	168.5	41.9	1136	4	AY609582	AY609582 Sus scrof
	39	168	41.8	504	3	BM504973	BM504973 ig88g06.y
	40	168	41.8	581	3	BM353974	BM353974 ig58g08.y
	41	168	41.8	582	6	CF249458	CF249458 it73f12.y
	42	168	41.8	611	6	CA773399	CA773399 im63f02.y
	43	168	41.8	618	5	BQ778644	BQ778644 il30g12.y
	44	168	41.8	622	2	BE969691	BE969691 601679632
	45	168	41.8	625	3	BI438227	BI438227 ic19e09.y

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 3, 2006, 23:30:21 ; Search time 163 Seconds
(without alignments)
199.943 Million cell updates/sec

Title: US-10-716-326-4
Perfect score: 402
Sequence: 1 MKIILWLCVFGLFLATLFPI.....SYLEGQAAKEFIAWLVKGRG 78

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	402	100.0	78	4	US-10-215-272-4	Sequence 4, Appli
2	402	100.0	78	4	US-10-716-326-4	Sequence 4, Appli
3	402	100.0	78	5	US-10-715-976-4	Sequence 4, Appli
4	364	90.5	80	4	US-10-215-272-20	Sequence 20, Appl
5	364	90.5	80	4	US-10-716-326-20	Sequence 20, Appl
6	364	90.5	80	5	US-10-715-976-20	Sequence 20, Appl
7	325.5	81.0	77	4	US-10-215-272-6	Sequence 6, Appli
8	325.5	81.0	77	4	US-10-716-326-6	Sequence 6, Appli
9	325.5	81.0	77	5	US-10-715-976-6	Sequence 6, Appli
10	324	80.6	87	5	US-10-775-180-202	Sequence 202, App
11	324	80.6	87	5	US-10-775-180-203	Sequence 203, App

12	324	80.6	87	5	US-10-775-204-635	Sequence 635, App
13	324	80.6	87	5	US-10-775-204-636	Sequence 636, App
14	198	49.3	64	4	US-10-360-101-228	Sequence 228, App
15	195	48.5	84	4	US-10-215-272-10	Sequence 10, Appl
16	195	48.5	84	4	US-10-716-326-10	Sequence 10, Appl
17	195	48.5	84	5	US-10-715-976-10	Sequence 10, Appl
18	192	47.8	82	4	US-10-215-272-8	Sequence 8, Appli
19	192	47.8	82	4	US-10-716-326-8	Sequence 8, Appli
20	192	47.8	82	5	US-10-715-976-8	Sequence 8, Appli
21	184.5	45.9	77	4	US-10-215-272-18	Sequence 18, Appl
22	184.5	45.9	77	4	US-10-716-326-18	Sequence 18, Appl
23	184.5	45.9	77	5	US-10-715-976-18	Sequence 18, Appl
24	176.5	43.9	77	5	US-10-775-180-686	Sequence 686, App
25	176.5	43.9	77	5	US-10-775-204-1791	Sequence 1791, Ap
26	176.5	43.9	662	5	US-10-775-180-611	Sequence 611, App
27	176.5	43.9	662	5	US-10-775-204-1623	Sequence 1623, Ap
28	176	43.8	654	5	US-10-775-180-574	Sequence 574, App
29	176	43.8	654	5	US-10-775-204-1559	Sequence 1559, Ap
30	172.5	42.9	89	5	US-10-775-180-449	Sequence 449, App
31	172.5	42.9	89	5	US-10-775-204-1282	Sequence 1282, Ap
32	172.5	42.9	673	5	US-10-775-180-834	Sequence 834, App
33	172.5	42.9	673	5	US-10-775-204-2170	Sequence 2170, Ap
34	172.5	42.9	674	5	US-10-775-180-447	Sequence 447, App
35	172.5	42.9	674	5	US-10-775-204-1280	Sequence 1280, Ap
36	172	42.8	663	5	US-10-775-180-600	Sequence 600, App
37	172	42.8	663	5	US-10-775-204-1609	Sequence 1609, Ap
38	171.5	42.7	60	5	US-10-775-180-835	Sequence 835, App
39	171.5	42.7	60	5	US-10-775-204-2180	Sequence 2180, Ap
40	171.5	42.7	83	5	US-10-775-180-684	Sequence 684, App
41	171.5	42.7	83	5	US-10-775-204-1789	Sequence 1789, Ap
42	171.5	42.7	145	5	US-10-775-180-685	Sequence 685, App
43	171.5	42.7	145	5	US-10-775-204-1790	Sequence 1790, Ap
44	171.5	42.7	664	5	US-10-775-180-598	Sequence 598, App
45	171.5	42.7	664	5	US-10-775-204-1607	Sequence 1607, Ap

ALIGNMENTS

RESULT 1
US-10-215-272-4
; Sequence 4, Application US/10215272
; Publication No. US20040002468A1
; GENERAL INFORMATION:
; APPLICANT: Genzyme Corporation
; APPLICANT: Wadsworth, Samuel C.
; APPLICANT: Armentano, Donna
; APPLICANT: Gregory, Richard J.
; APPLICANT: Parsons, Geoffrey
; TITLE OF INVENTION: Methods of Treating Diabetes and Other
; TITLE OF INVENTION: Blood Sugar Disorders
; FILE REFERENCE: 2478.2019002 PCT
; CURRENT APPLICATION NUMBER: US/10/215,272
; CURRENT FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: US 60/310,982
; PRIOR FILING DATE: 2001-08-08
; NUMBER OF SEQ ID NOS: 54

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 78
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Amino acid sequence of Exendin-4.GLP-1Gly8
US-10-215-272-4

  Query Match             100.0%;  Score 402;  DB 4;  Length 78;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-39;
  Matches   78;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKIILWLCVFGLFLATLFPISWQMPVESGLSSEDSASSESFAKRIKRHGEGTFTSDVSSY 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKIILWLCVFGLFLATLFPISWQMPVESGLSSEDSASSESFAKRIKRHGEGTFTSDVSSY 60

Qy         61 LEGQAAKEFIAWLVKGRG 78
              ||||||||||||||||||
Db         61 LEGQAAKEFIAWLVKGRG 78


RESULT 2
US-10-716-326-4
; Sequence 4, Application US/10716326
; Publication No. US20040143104A1
; GENERAL INFORMATION:
;  APPLICANT: Genzyme Corporation
;  APPLICANT:  Wadsworth, Samuel
;  APPLICANT:  Armentano, Donna
;  APPLICANT:  Gregory, Richard J.
;  APPLICANT:  Parsons, Geoffrey
;  TITLE OF INVENTION: Methods of Treating Diabetes and Other Blood Sugar
Disorders
;  FILE REFERENCE: 5062CIP
;  CURRENT APPLICATION NUMBER: US/10/716,326
;  CURRENT FILING DATE:  2003-11-17
;  PRIOR APPLICATION NUMBER: US 10/215,272
;  PRIOR FILING DATE: 2002-08-07
;  PRIOR APPLICATION NUMBER: US 60/310,982
;  PRIOR FILING DATE: 2001-08-08
;  NUMBER OF SEQ ID NOS: 54
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
;   LENGTH: 78
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Amino acid sequence of Exendin-4.GLP-1Gly8
US-10-716-326-4

  Query Match             100.0%;  Score 402;  DB 4;  Length 78;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-39;
  Matches   78;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKIILWLCVFGLFLATLFPISWQMPVESGLSSEDSASSESFAKRIKRHGEGTFTSDVSSY 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          1 MKIILWLCVFGLFLATLFPISWQMPVESGLSSEDSASSESFAKRIKRHGEGTFTSDVSSY 60

Qy         61 LEGQAAKEFIAWLVKGRG 78
              ||||||||||||||||||
Db         61 LEGQAAKEFIAWLVKGRG 78


RESULT 3
US-10-715-976-4
; Sequence 4, Application US/10715976
; Publication No. US20050107318A1
; GENERAL INFORMATION:
;  APPLICANT: Genzyme Corporation
;  APPLICANT:  Wadsworth, Samuel
;  APPLICANT:  Armentano, Donna
;  APPLICANT:  Gregory, Richard J.
;  APPLICANT:  Parsons, Geoffrey
;  TITLE OF INVENTION: Methods of Treating Diabetes and Other Blood Sugar
Disorders
;  FILE REFERENCE: 5121
;  CURRENT APPLICATION NUMBER: US/10/715,976
;  CURRENT FILING DATE:  2003-11-17
;  NUMBER OF SEQ ID NOS: 54
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
;   LENGTH: 78
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Amino acid sequence of Exendin-4.GLP-1Gly8
US-10-715-976-4

  Query Match             100.0%;  Score 402;  DB 5;  Length 78;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-39;
  Matches   78;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKIILWLCVFGLFLATLFPISWQMPVESGLSSEDSASSESFAKRIKRHGEGTFTSDVSSY 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKIILWLCVFGLFLATLFPISWQMPVESGLSSEDSASSESFAKRIKRHGEGTFTSDVSSY 60

Qy         61 LEGQAAKEFIAWLVKGRG 78
              ||||||||||||||||||
Db         61 LEGQAAKEFIAWLVKGRG 78


RESULT 4
US-10-215-272-20
; Sequence 20, Application US/10215272
; Publication No. US20040002468A1
; GENERAL INFORMATION:
;  APPLICANT: Genzyme Corporation
;  APPLICANT:  Wadsworth, Samuel C.
;  APPLICANT:  Armentano, Donna
;  APPLICANT:  Gregory, Richard J.
;  APPLICANT:  Parsons, Geoffrey
;  TITLE OF INVENTION: Methods of Treating Diabetes and Other
;  TITLE OF INVENTION:  Blood Sugar Disorders
```